Published in final edited form as:

Horm Behav. 2012 January ; 61(1): 100–107. doi:10.1016/j.yhbeh.2011.10.009.

Compatibility drives female preference and reproductive success in the monogamous California mouse (*Peromyscus californicus*) more strongly than male testosterone measures

Erin D. Gleason^{a,*}, Mary A. Holschbach^a, and Catherine A. Marler^{a,b}

^aDepartment of Psychology University of Wisconsin-Madison, USA

^bDepartment of Zoology, University of Wisconsin-Madison, USA

Abstract

Female assessment of male attractiveness and how preferred qualities impact reproductive success is central to the study of mate choice. Male attractiveness may depend on traits beneficial to the reproductive success (RS) of any female, termed ‘universal quality’, and/or on behavioral and biological interactions between potential mates that reflect ‘compatibility’. The steroid hormone testosterone (T) often underlies male attractiveness in rodents and is associated with enhanced paternal care in the monogamous and biparental California mouse (*Peromyscus californicus*). We hypothesized that (1) T-characteristics are universally attractive to female California mice and that (2) if reproductive success is higher for females mated with preferred males, then females mated with males preferred by other females will also have higher reproductive success. Alternatively, we speculated that pair compatibility, based on emergent pair qualities, is important for a species with coordinated offspring care. We assessed individual T-characteristics in three ways: (1) T-response to GnRH challenges (2) baseline T-level and (3) T-response to a female. Testosterone-response did not predict female preference, but females spent more time investigating males with higher baseline T (accounting for only 9.6% of the variation in investigation time). None of the T-measures was associated with RS. Females paired with males they preferred produced litters more quickly and had higher RS than females paired with their non-preferred males. Naïve females who did not undergo preference tests had equivalent RS regardless of whether their mate was preferred or non-preferred by another female. These data suggest that higher male T elicits investigation, but female preference in the California mouse is more strongly linked with compatibility because individual preference was a better predictor of RS than any T measure.

Keywords

Monogamy; *Peromyscus*; Paternal; Testosterone; Preference; Compatibility

*Corresponding author at: Tufts University Cummings School of Veterinary Medicine, 200 Westboro Road, North Grafton, MA 01536, USA. erin.gleason@tufts.edu (E.D. Gleason).

Introduction

Mate choice is one of the most pivotal processes in biology, as these decisions can influence which genes are passed to future generations to ultimately shape the physiology and behavior of a species. Strategies for mate choice vary greatly among individuals, species and life history stages, typically reflecting a desire to maximize benefits to oneself or one's offspring while limiting the negative outcomes associated with a poor choice (Trivers, 1972). Although active mate choice is unlikely to revolve around any single factor (Wagner, 1998), there are some distinctions that can be made regarding what drives female preference for certain males. Female preference for certain males could indicate that these males have traits beneficial to the reproductive success (RS) of any female who chooses them, which we will term 'universal quality' (Andersson, 1982, 1994; von Schantz et al., 1989). Alternatively, female preferences could depend on the level of 'compatibility' between the male and female, reflecting the quality of behavioral or biological interactions between specific individuals that in turn determine pair RS (Ryan and Altmann, 2001). These two models of male attractiveness are not mutually exclusive and may be weighted differently depending on a number of factors including species social organization and availability of mates.

The breeding system of a species is one factor that may underscore the importance of active mate choice. The monogamous and biparental California mouse, *Peromyscus californicus*, is a genetically monogamous species in which individuals typically have only one mate in a lifetime (Ribble, 1991, 1992). In addition, offspring survival relies on the presence and care of both the mother and father (Gubernick and Teferi, 2000). Taken together, these data suggest that whether a California mouse successfully produces offspring is significantly dependent on the onetime choice of a mate.

Testosterone (T) is one biological factor that has been shown to support universal attractiveness of male odors in rodents (Ferkin et al., 1992; Taylor et al., 1982), and may be of particular significance for mate choice in the California mouse given that it is associated with critical male social behaviors in this species; recently, we reported that male T-response to courtship interactions predicts future paternal behavior, specifically in the amount of pup huddling and grooming performed when the female is temporarily absent (Gleason and Marler, 2010). Moreover, these findings are consistent with a general promotion of paternal behavior by T in the California mouse (Trainor and Marler, 2001, 2002). Indeed, examples of female preference for males who signal their paternal competence are found in other taxa (Forsgren, 1997; Lindström et al., 2006; Ostlund and Ahnesjö, 1998), although this relationship has yet to be shown in mammals. Our goal was to test whether male quality is primarily associated with ability to express high levels of T, specifically in a monogamous mammal. Given the importance of paternal care and its dependence on T in the California mouse (Trainor and Marler, 2001, 2002), in the present study we tested the hypothesis that female California mice prefer males who mount higher T-responses to courtship interactions, or have higher baseline T. In showing mating preferences for higher-T males, female California mice may be able to secure higher levels of paternal behavior, which would be universally beneficial to the RS of any female who chooses them.

Alternatively, mate preference in female California mice could reflect a focus on compatibility between mates, in which case mate preference and/or RS would be more closely associated with individual preference rather than a consistent male trait such as T-response. Few studies have considered how male qualities are differentially weighted when there are multiple factors to which a female can attend; that is, we accounted for the possibility that females would show mate preferences, but not for T-characteristics. Thus, we also tested the alternative hypothesis that female preference and male attractiveness are based on pair quality derived from compatibility, not universal quality.

To assess whether female preference is associated with male T characteristics, we first used gonadotropin releasing hormone (GnRH) challenges in combination with blood sampling to determine whether male T-measures or response characteristics are consistent within individual males. We then conducted preference tests in which we allowed female California mice to choose between males with high or low T-responses to GnRH challenges, and compared RS of breeding pairs comprised of a female paired with a high or low T-response male. In addition, we collected blood samples from males 1 day prior to and 1-h after preference testing to compare T-response to GnRH challenge with T-response to a female. To test the alternative hypothesis that mate choice is driven by compatibility, meaning that preferences are based on emergent pair qualities that cannot be predicted from either individual alone, we measured the relative impacts of preference and T on RS. We compared the RS of pairs formed after preference testing in the following four categories: 1) females that underwent preference tests and were paired with their preferred (P) males, 2) females that underwent preference tests and were paired with their non-preferred (NP) males, 3) naïve females that did not experience a preference test and were paired with a P male (preferred by another female in preference tests) and 4) naïve females that did not experience a preference test and were paired with a NP male (male that was not preferred by another female in preference tests). By comparing these four groups, we were able to simultaneously investigate the impact of both male T-characteristics (a putative universal benefit) and preference (which may or may not be T-based) on RS in the California mouse. We expected that if preferences are based on an assessment of either biological or behavioral compatibility between the male and female, then the strongest predictor of pair RS would be whether females are paired with males that *they* preferred during preference tests.

Methods

Animals

Forty-six male California mice between 5 and 6 months old were chosen at random from our breeding colony. Twelve of the 46 animals initially served as saline controls, and then again as subjects for GnRH challenges after a two-week rest period. Animals were housed in standard laboratory cages (48.3-cm-long×26.7-cm-wide×15.6-cm-high) with one or two same-sex cage mates during GnRH challenges and prior to preference testing, and later as breeding pairs. Purina 5015 mouse chow and drinking water were provided ad libitum. The testing room was maintained at 25°C on a 14:10 light/dark cycle with all behavioral observations occurring under dim red light. Animals were cared for in accordance with the *National Institutes of Health Guide for the Care and Use of Laboratory Animals*, and all

procedures were approved by the University of Wisconsin-Madison IACUC. For clarification, a timeline of study events can be found in Chart 1.

GnRH challenges

Exogenously administered GnRH triggers the hypothalamic–pituitary gonadal (HPG) axis, causing an endogenous cascade of endocrine signals that ultimately results in T secretion from the testes. At low doses, the change in T is correlated with the amount of GnRH administered. At the high doses used during a GnRH challenge, however, it is possible to measure the maximum response that a male is capable of producing (Lacombe et al., 1991). By administering a standardized, high dose of GnRH to each animal, resultant plasma T-levels reflect individual differences in the ability to rapidly produce T (Jawor et al., 2006). GnRH challenges were conducted in two rounds, one in October of 2009 ($n=26$ animals) and the second in January of 2010 ($n=24$ animals). Challenges occurred 30 to 60-min after lights-out, under dim red light. Once per week for three consecutive weeks, each male was tested as follows: mice were transported in their home cages from the colony to a laboratory room directly across the hall on a wheeled cart covered to prevent exposure to hallway lighting. Each male was lightly anesthetized using Isoflurane anesthesia and 60 μ L of blood were drawn from the left eye with a heparinized capillary tube. Capillary tubes were immediately placed in a 5-mL microcentrifuge tube on ice. Light pressure was applied to the eye to ensure that bleeding had stopped, and the mouse was immediately injected with 50.0 ng of i.p. GnRH (Sigma, catalog # L7134) dissolved in 1.0 cm³ of 0.9% sterile saline; the appropriate GnRH dose and time-course of sampling were both determined in pilot studies and based on doses reported in the literature (Gábor et al., 1998; Jawor et al., 2006; Millesi et al., 2002). The mouse was weighed, returned to his home cage and transported back to the colony to minimize the stress associated with the scent of blood sampling from other animals in the laboratory. 45 min after GnRH injections, animals were returned to the laboratory, lightly anesthetized and another 60 μ L of blood was collected from the right eye. Blood samples were centrifuged at 10,000 rpm (Eppendorf, model 5417R) for 10-min, and plasma was stored at -80°C until assays were conducted. Analysis of T-samples for both GnRH challenges and mate preference tests presented in this paper was performed by the Wisconsin National Primate Research Center. Steroids were extracted twice with ethyl ether, and separated using celite chromatography. Testosterone was analyzed using enzyme immunoassay (T antibody R156, UC-Davis diluted to 1:35,000, validated for California mice by Bester-Meredith and Marler, 2001; Trainor and Marler, 2001; Davis and Marler, 2003). The inter- and intra-assay coefficients of variation were 15.2% and 3.5%, respectively ($n=7$ plates). After removing data points from individuals whose baseline T measures indicated a naturally-occurring T-pulse had occurred just before testing (Coquelin and Desjardins, 1982); (one T-measure $>400\%$ of either of the two other baseline T-measures from the same individual; challenge 1=6 animals removed, challenge 2=5 animals and challenge 3=6 animals), final sample size for T-response to GnRH in each of three trials was $n=44$, $n=45$, and $n=44$, respectively. Note that an animal removed from an individual GnRH challenge could be represented in the other two challenges.

Analysis of T-response to GnRH

T-response was calculated by subtracting baseline T from final T (Vickers, 2001). Pearson correlations were used to confirm that individuals were consistent in their T-responses across GnRH trials (see results), at which time a mean T-response was calculated for each individual. For three individuals, mean T-response was calculated using data from two GnRH challenges. The final sample size was 45 animals. Mean T-responses to GnRH challenge were used to classify the animals as “high” or “low” responders for preference testing, using a median split of the sampled population. Within each category, males were rank-ordered by their mean T-responses, and dyads of stimulus males were formed by matching males with the same rank from each category (e.g., males with a rank of “1” from each category, representing the male who was the highest of the “high” responders and the male who was the highest of the “low” responders). The purpose of these groupings was to increase the probability that stimulus males had differing physiological abilities to mount a T-response during mate preference testing. Using this procedure, pairs of stimulus males were separated by a minimum of one standard deviation of the mean T-response to GnRH (see Results section).

Mate preference tests

The mate preference testing apparatus consisted of a polycarbonate testing cage (91-cm long×46-cm wide×43-cm high) divided into three equal chambers 46-cm long×33.33-cm wide (Fig. 1). The central chamber was an empty, neutral area and the left and right chambers (distal) were further partitioned into 30.66-cm long holding compartments for stimulus males. These holding compartments were separated from the distal chambers by 1”×1” gage wire mesh. Opaque walls separated the three main chambers. Entrances to either distal chamber were located at the front of the apparatus and could be individually blocked by sliding an opaque divider into place. Clear polycarbonate lids with drilled air holes were placed on the top of the entire apparatus to prevent animals from escaping during testing.

Baseline blood samples were collected from males 1 day prior to preference testing, 1–2 h after lights-out. The day before testing, males spent 30-min in one of the holding chambers of the preference apparatus to acclimate. On the morning of preference testing, we performed vaginal lavage on a group of 8–12 females to identify estrus by the exclusive presence of cornified epithelial cells in vaginal lavage (Caligioni, 2009; Dewsbury, 1974); once we had found two females that were in estrus, we also confirmed that the females shared no ancestry with either stimulus male for a minimum of two generations. Next we placed one of the estrous females in the testing apparatus for a 30-min habituation period, during which she had access to both the central chamber and empty distal chambers. Immediately before each trial, the divider to each chamber was replaced, enclosing the female in the central chamber. The high-T and low-T stimulus males were then placed in each of the distal chambers. The placement of low/high T males in either the right or left chamber was counterbalanced across trials using a letter code; an undergraduate assistant assigned one of two letters (A or B) representing T-response characteristics to each mouse, and each trial consisted of one mouse in category “A” and one mouse in category “B”. Because the experimenter did not know which category each letter represented, she was

blind to group identity during preference testing (either high/low T). To begin the preference test, the sliding door to the left chamber was raised and then lowered once the female had voluntarily entered the chamber. After a period of 5-min, the door was raised and lowered again after the female had exited to the neutral chamber. The process was then repeated for the right chamber, concluding with the return of the female to the neutral chamber. After a 30-min waiting period to allow the male adequate time to mount a T-response (Trainor et al., 2004), the dividers between the neutral and distal chambers were simultaneously raised. Preference tests lasted for 15-min, and determination of preference excluded time spent in the neutral chamber. Trials were recorded on DVD for backup, but the duration of time a female spent in each of the left and right chambers containing a stimulus male was scored in real-time. We collected blood samples from the males immediately following preference testing (45–50 min post-initial exposure to the female), and returned them to their home cages.

We used a baseline criterion to establish that females were in fact investigating the stimulus males. In all 23 preference tests, females spent at least 50% (Drickamer et al., 2000) of the 15-min trial interacting with either of the males, indicating motivation to investigate ($M=79.78\%$ of total trial time). To establish which male the female preferred, we assumed that the male with which the female spent a greater duration of that time would accurately reflect her mate preference (minimum duration of time a female spent with a male categorized as “preferred” was 55% of the total time spent with either male; time spent with either male ranged from 13.9%–86.1%, and >60% of time spent with one male was considered to represent a strong preference). With the exception of one trial that was excluded from analysis because the female spent equal time with both males, we included the remaining 22 trials in our analysis. 2 h after mate preference testing, animals were randomly assigned to one of four pairing conditions; naïve pairings were formed using the second estrous female that we had identified that same morning. Sample size for choosing females with their preferred male was $n=12$, choosing females with their non-preferred male was $n=11$, naïve females with a preferred male was $n=11$, and naïve females with a non-preferred male was $n=12$.

Reproductive success

In the field, a majority of California mice breed for only one season, and the maximum documented number of litters produced by any pair is four (Ribble, 1992). To estimate RS, we recorded breeding data from pairs for 135 days post-pairing, the minimum number of days required for any pair in the study to produce four litters. The number of litters and pups produced for each pair, as well as the mass of each pup born, were collected continuously and also summarized at two time-points during the study; 68-days post-pairing (halfway) and 135 days (the end of the study) to reflect the time required for pairs to produce the maximum of four litters (Ribble, 1992). Analyzing data at the halfway point of 68-days allowed us to maximize the number of pairs included in the analysis (for six pairs, we were missing weight data on the second litter). The number of days between pairing and the birth of the first litter for each pair was also recorded.

Statistical analysis

All variables were checked for normality using q–q plots before analysis and a natural-log transformation (Osborne, 2002) was applied to baseline T-measures collected on the day before mate preference testing. We used independent-samples t-tests to assess group differences in two-item categorical variables with continuous outcomes. Chi-square tests were used to determine whether females tended to choose males from any category (e.g., “in how many cases did a female choose the higher-T male of the two males with which she was presented?”). Pearson correlations were used to test for relationships between continuous behavioral and endocrine variables, with the Benjamini–Hochberg correction used for false discovery by multiple comparisons (Benjamini and Hochberg, 1995).

Results

T-response to GnRH challenge

An independent samples *t*-test confirmed that a 50-ng injection of GnRH was effective in increasing T-levels compared to a saline injection (corrected $t_{36.37} = -3.277$, $p = 0.002$) (Fig. 1). We collected data on 46 individuals (T-response, $M = 818.16$ pg/mL, $SD = 505.33$ pg/mL). Because we gathered data in two batches separated in time by 4 months, we also used independent samples *t*-tests to confirm that the distribution of T-response to GnRH challenge did not significantly differ between these groups of animals ($t_{44} = 1.17$, $p = 0.25$). In addition, there was no relationship between any measure of post-GnRH challenge T and body weight (*p*-values ranged from 0.2 to 0.9).

Testosterone-response to GnRH challenge was highly correlated within individuals across all trials. Testosterone-response in trial one was positively correlated with T-response in trials two ($r = 0.62$, $p < 0.001$) and three ($r = 0.67$, $p < 0.001$), and T-responses in trials two and three were also correlated ($r = 0.48$, $p < 0.001$). However, we unexpectedly found that the magnitude of individual T-response was not consistent over time, but on average amplified in the third trial according to the analysis below. Mean T-response to GnRH challenges across the three trials was compared with repeated-measures ANCOVA using baseline T as a covariate and change in T as the dependent measure. Mauchly’s test indicated that a violation of sphericity had occurred ($\chi^2(2) = 7.15$, $p = 0.028$), so degrees of freedom were corrected using the Huynh–Feldt model estimates of sphericity ($\epsilon = 0.96$). The overall model showed that there was no significant main effect of trial ($F_{2,74} = 0.56$, $p = 0.57$). Tukey Honestly Significant Difference post-hoc tests revealed that while within-subjects T-responses in trials one ($M = 699.56$, $SD = 510.38$) and two ($M = 733.62$, $SD = 683.34$) were statistically equivalent, T-responses in trials one and three ($M = 968.80$, $SD = 803.87$) differed at the $p < 0.05$ level. Amplification of T-responses in GnRH trial three was driven by 24 animals whose baseline and final T-values had risen as compared to earlier tests, while T-values for the remaining 22 animals did not change across trials. Although this phenomenon will not be discussed in this paper, it is an intriguing finding that warrants further study.

Mate preference and testosterone

T-response to GnRH challenge and mate preference tests—The duration of time that females spent with males during mate preference tests was not correlated with any

aspect of T-response to GnRH challenge, including average T-response across the three trials, baseline or final T measures from any single trial, or T-response to any one of the three GnRH challenges (all p -values > 0.17). When comparing preferred versus non-preferred males, there were non-significant trends for preferred males to have higher baseline ($t_{42} = -2.10$, $p = 0.051$) and final T-measures ($t_{41} = -1.89$, $p = 0.065$) during the third GnRH challenge, but no other group differences (p -values > 0.42). Of 22 trials, in 14 cases females preferred males who had experienced amplification to GnRH challenge, although this only reached trend-level significance ($\chi^2 = 2.87$, $p = 0.091$). When examining just 11 trials (22 males) during which females had shown the strongest preferences ($> 60\%$ time with one male), the same pattern of results held, and the trend for females to prefer males with higher baseline T on the third GnRH challenge became statistically significant ($t_{20} = -2.58$, $p = 0.018$). This is an interesting finding that warrants further research, but because the average for the three baseline T measures during GnRH challenges was not significant even for those expressing the strongest preferences, this is not viewed as contributing to female preference under the conditions of the current experiment.

Natural T-response and mate preference—We also tested whether natural T-response to an estrous female during mate preference testing (as opposed to T-response to a GnRH injection) impacted female preference. The duration of time females spent with males during mate preference tests was not correlated with final T or T-response measures (p -values 0.82 and 0.27, respectively), but there was a significant positive correlation between the duration of time a female spent with a male and his baseline T level (measured on the day prior to preference testing; $r = 0.31$, $p = 0.043$, Fig. 2). The amount of variation in time spent with a male that can be explained by baseline T is equal to R^2 , or 9.6%. Given that this correlation reflected the entire population of animals used in preference tests, we also conducted a set of Chi-square tests comparing the baseline T measures between the two specific males from which a female chose in the 11 trials representing stronger ($> 60\%$) preferences. In 10 of 11 trials the female preferred the male who had higher baseline T ($\chi^2 = 13.75$, $p < 0.001$). Distribution of preferred and non-preferred males between higher or lower final T, and higher and lower T-responses, was that which would be expected by chance (p -values = 0.67 and 0.16, respectively). Thus, higher baseline T, not final T or T-response, appears to be the only T feature that influences female preference.

Plasma samples from GnRH challenges and mate preference tests were analyzed in separate assays, so hormone concentrations are not directly comparable between time points. However, we did use Spearman rank-order correlations to determine whether a relative rank was maintained within the population across the different types of tests, and found that average T-response to GnRH challenge was unrelated to T-response to females ($p = 0.44$). We also examined the distribution of individuals who were in the highest 50% of responders to GnRH and also within the highest 50% of responders to a female, and found no relationship ($p = 0.75$). Thus, T-response to GnRH was not predictive of T-response to a female.

Reproductive success and mate preference

At the conclusion of the study (135 days paired), pairs produced an average of 2.29 litters ($SD = 0.96$), ranging from 0 to 4 litters. The mean number of pups born was 5.36 ($SD = 2.87$,

range 0–12). Four pairs produced zero litters, and one female died giving birth. To assess reproductive success, we constructed a composite RS variable by calculating the mean of the 68-day z-scores of three highly correlated variables (number of litters born, number of pups born, and total birth weight of pups, p -values<0.001). The birth weight of pups was highly correlated with weight at weaning (p <0.001), so we used only birth weight in our measure of RS. Final sample size for number of days to produce a litter was 41 because four pairs never produced a litter, and 45 pairs for RS at 68 days.

Reproductive success by pair type for females in preference tests—Analysis of pair type considered two groups, each comprised of females that had undergone mate preference testing. The first group consisted of females paired with their non-preferred (NP) male, and the second group was females paired with their preferred (P) male. We found that females paired with their P male produced litters faster ($M=46.82$ days, $SD=13.33$, $n=12$) than females paired with their NP male ($M=67.7$ days $SD=21.65$, $n=11$), (corrected $t_{14.70} = -2.63$, $p=0.019$) (Fig. 3). At the 68-day point, females paired with their P male also had higher RS scores ($M=0.54$ $SD=0.90$) than females paired with their NP male ($M=-0.52$, $SD=0.86$) ($t_{21}=2.87$, $p=0.009$) (Fig. 4), which is likely due to group differences in latency to produce a litter.

Reproductive success by pair type for naïve females (no experience with preference tests)—Comparing P and NP males paired with naïve females, there were no group differences in latency to produce a first litter or RS at 68 days (p -values=0.07, 0.09 respectively); note that if anything, trends were toward P males having *lower* RS and a *greater* number of days to produce a litter than NP males (Fig. 4). This pattern may also be interesting to investigate in the future.

Reproductive success and testosterone

Reproductive success and T-response to GnRH challenge—Neither mean T-response to GnRH challenge nor any baseline, final, or change measures from individual GnRH challenges were related to latency to produce a litter or RS at 68 days (p -values>0.20).

Reproductive success and natural T-response to mate preference—Baseline T, final T, and change T in response to mate preference trials were not correlated with latency to produce a litter or RS at 68 days (p -values>0.25).

Discussion

Neither average T-response to GnRH challenge nor T-response to a female was associated with female preference, or with enhanced RS. Overall, females spent more time investigating males with higher baseline T measured just prior to mate preference tests, but this did not ultimately predict female preference. Female preference was associated with RS; California mouse females paired with males they preferred produced litters more quickly and had higher RS scores than pairs consisting of females paired with their non-preferred males. As described below, higher baseline T may contribute to female preference, but our

results indicate that individual preference, and not T characteristics, ultimately predicts RS in this species.

Our finding that female mate preference is predictive of RS is consistent with other work in house mice confirming a reproductive advantage to females who are permitted to exercise mating preferences in a polygynous species (Drickamer et al., 2000; Drickamer et al., 2003), and with results from another monogamous *Peromyscus* species showing that male mating preferences are similarly based on compatibility (Ryan and Altmann, 2001). It is noteworthy that similar patterns of reproductive success on the basis of either male or female preference are observed in both polygynous and monogamous species. In house mice (Drickamer et al., 2000; Drickamer et al., 2003), matings in which either females had preferred males or males and females had mutually preferred each other, offspring viability was higher as assessed by several measures. For example, adult sons from P matings were socially dominant to males from NP matings and adult offspring of both sexes built superior nests. In field enclosures, fewer NP offspring survived to 60 days than offspring from P matings. The authors note that although the specific reasons for these differences were not examined, the findings are consistent with an enhanced immunocompetence model in which offspring from P matings have greater resistance to pathogens. Thus, compatibility appears to be important for RS in rodents regardless of mating system, but these similar findings may result from different types of compatibility between the male and female; for example, biological compatibility leading to higher immunocompetence in offspring clearly could be important for species, such as the house mouse, in which there is little or no paternal care. In a monogamous species, the assessment of compatibility may shift toward a focus on behavior, and particularly behavior relating to male parental care. Nonetheless, it seems that the best mate for one individual may not be the best for another in both polygynous and monogamous mammals. Field studies of California mouse populations have shown that one of the most important predictors of lifetime reproductive success is timing of first reproduction, because litter size and inter-birth interval remain consistent across subsequent litters (Ribble, 1992). Higher lifetime RS for pairs that begin reproducing more quickly indicates that phenotypic selection is likely to occur on this aspect of breeding (Ribble, 1992), lending support to the importance of our finding that preference dictates when pairs first produce a litter. Our observed group differences in time to produce a litter and RS were most likely the result of behavioral mechanisms to delay mating in non-preferred pairings. For example, females may have resisted advances by NP males (Gleason and Marler, 2010) or males may also have recognized a poor fit and postponed mating attempts. Either individually or in concert, these behaviors would certainly delay copulation and pair bond formation. While our NP pairings eventually resulted in offspring in the laboratory, under field conditions these delays in mating would likely allow animals to reject a pairing and continue to look for a more suitable mate. Finally, we note that the presence of genetic monogamy in this species suggests that males would also exercise mating preferences rather than mate indiscriminately. Although our focus was on female mating preferences in the current paper, and due to faster litter production for pairs in which females had shown a preference we can infer that at least the minimum mating criteria for males were met, the influence of male preference on RS should also be examined.

Female rodents of several species prefer the urinary odors of intact versus castrated adult males, indicating that gonadal hormones influence the attractiveness of these odors (Ferkin et al., 1992; Ferkin and Johnston, 1993; Taylor et al., 1983). Thus, the observed increase in female investigation of higher baseline-T males could reflect sexual attraction to odor cues that indicate the presence of a reproductively active male. From a signaling perspective, female interest in these odors helps ensure that animals are drawn together when the male is reproductively mature and able to inseminate the female. Females further discriminate among intact males on the basis of variations in their T-levels in some species, generally preferring to spend more time investigating odors from males with higher plasma T (Ferkin et al., 1994; Gottreich et al., 2000). Thus, higher levels of androgens may simply result in a greater production of the components of the chemosignal that elicit female investigation, but a combination of additional cues ultimately determines whether or not a male is preferred.

Our results suggest that RS is the result of emergent pair qualities that are not predicted by the quality of the male alone (Spoon et al., 2006); females paired with their preferred males from the preference tests experienced a higher estimate of reproductive success that was not experienced when females were mated to males preferred by other females. Moreover, T levels themselves did not directly impact reproductive success. If female mate choice in the California mouse was based solely on a trait or traits universally valuable to all females, we would have expected to see enhanced RS in naïve females paired with males preferred by a choosing female. In general, emergent properties of pair compatibility have been relatively understudied. If one assumes a biological definition of compatibility, one likely candidate is the major histocompatibility complex (MHC), a group of genes that support immune functioning and are key predictors of mate choice in house mice (Yamazaki et al., 1976). On a behavioral level, compatibility may be contingent on the interactions between the male and female. For example, in cockatiel pairs with higher measures of behavioral compatibility (as measured by amicable behaviors such as coordinated behavior, heightened behavioral synchrony, allopreening and sexual behavior) rear more chicks to independence (Spoon et al., 2004, 2006). In a biparental species, many behaviors need to be coordinated such that the pair can manage offspring care, foraging, adverse environmental conditions, predators, and the needs of the mate. Finally, we must consider that in some cases behavioral compatibility may arise from universal quality; for example, males that are generally more responsive to the needs of their mates may have higher measures of pair compatibility, but this trait can be considered universal if it is independent on the particular female to which he is paired. Although we treated universal quality and compatibility separately in this paper, their interaction presents the opportunity for intriguing future research.

Differences in social organization, including breeding system, are likely to be important considerations regarding the use of GnRH challenge in future studies. GnRH challenge is a relatively new technique to be used in behavioral endocrinology, and we are still in the process of understanding when and how it is best implemented. It seems clear that using GnRH challenges to understand behavior is most appropriate for species that show socially-induced modulation of T, which may occur in either aggressive or sexual contexts (Gleason et al., 2009). In the present study, we confirmed that T-response to GnRH challenge is highly consistent within individual male California mice, but contrary to predictions, these characteristics do not appear to be preferred, advantageous to RS, or even consistent with T-

response to a female; however, it is possible that such relationships would arise over multiple days or encounters with a female. Still, our findings follow a different pattern than in other species such as the dark-eyed junco, for which individual variations in T-response to GnRH predict variations in both behavior and fitness (Jawor et al., 2006; McGlothlin et al., 2008; McGlothlin et al., 2007; McGlothlin et al., 2005). In retrospect, species differences in reproductive behavior might explain why we did not find a similar relationship in the California mouse. Although the dark-eyed junco is socially monogamous, males seek and obtain extra-pair fertilizations, indicating that full activation of the HPG axis upon encountering a new female may be reproductively advantageous (Ketterson et al., 1997; Raouf et al., 1997). In contrast, a paired California mouse that encounters a novel female should not (Gubernick and Nordby, 1993; Ribble, 1991) engage with that female sexually, and even unpaired males may be more cautious in initiating sexual interactions than in a polygynous species. Because GnRH challenge bypasses higher-level (HPG) effects, its use here may have neglected important social modulators of T-response to females, specifically those drawn from interactions with a prospective mate. In support of a role for social modulation of T-response to female stimuli, one study in common marmosets (*Callithrix jacchus*) found that T-response to ovulatory odors was blocked in paired fathers, whereas paired males without infants showed the classic rise in T (Ziegler et al., 2005). Incidentally, the California mouse also mounts a T-response to winning an aggressive encounter (Fuxjager et al., 2010; Fuxjager et al., 2009; Gleason et al., 2009; Oyegbile and Marler, 2005). Whereas reproductive stimuli might require dampened responsiveness of the HPG axis, the appropriate behavioral response for a California mouse faced with a territorial intruder is most likely going to be aggression. We speculate that T-response to GnRH challenge would better predict individual differences in aggressive behavior and T-response to winning an aggressive encounter in this species.

In summary, we present evidence that RS and latency to begin producing litters in the California mouse depend, at least in part, on female perception of compatibility. The mechanism by which California mice evaluate compatibility is not known, but may relate to either immune (MHC) or behavioral compatibility, and would be an interesting direction for future research.

Acknowledgments

We thank E. Herget for assistance with the experiments. This research was supported by NSF DDIG 1010799 to E.D.G., NSF Grant IOS-0620042 and the Wisconsin Alumni Research Foundation to C.A.M. and NIH grant RR000167 (Wisconsin National Primate Research Center, University of Wisconsin-Madison).

References

- Andersson M. Female choice selects for extreme tail length in a widowbird. *Nature*. 1982:818–820.
- Andersson, M. *Sexual Selection*. Princeton University Press; Princeton: 1994.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Methodol*. 1995; 57:289–300.
- Bester-Meredith JK, Marler CA. Vasopressin and aggression in cross-fostered California mice (*Peromyscus californicus*) and white-footed mice (*Peromyscus leucopus*). *Horm Behav*. 2001; 40:51–64. [PubMed: 11467884]

- Caligioni CS. Assessing reproductive status/stages in mice. *Current Protocols in Neuroscience Appendix*. 2009; 4
- Coquelin A, Desjardins C. Luteinizing hormone and testosterone secretion in young and old male mice. *Am J Physiol*. 1982; 243:E257–E263. [PubMed: 7114252]
- Davis ES, Marler CA. The progesterone challenge: steroid hormone changes following a simulated territorial intrusion in female *Peromyscus californicus*. *Horm Behav*. 2003; 44(3):185–198. [PubMed: 14609541]
- Dewsbury DA. Copulatory behavior of California mice (*Peromyscus californicus*). *Brain Behav Evol*. 1974; 9:95–106. [PubMed: 4412636]
- Drickamer LC, Gowaty PA, Holmes CM. Free female mate choice in house mice affects reproductive success and offspring viability and performance. *Anim Behav*. 2000; 59:371–378. [PubMed: 10675259]
- Drickamer LC, Gowaty PA, Wagner DM. Free mutual mate preferences in house mice affect reproductive success and offspring performance. *Anim Behav*. 2003; 65:105–114.
- Ferkin MH, Johnston RE. Roles of gonadal hormones in control of five sexually attractive odors of meadow voles (*Microtus pennsylvanicus*). *Horm Behav*. 1993; 27:523–528. [PubMed: 8294120]
- Ferkin MH, Gorman MR, Zucker I. Influence of gonadal hormones on odours emitted by male meadow voles (*Microtus pennsylvanicus*). *J Reprod Fertil*. 1992; 95:729–736. [PubMed: 1404090]
- Ferkin MH, Sorokin ES, Renfro MW, Johnston RE. Attractiveness of male odors to females varies directly with plasma testosterone concentration in meadow voles. *Physiol Behav*. 1994; 55:347–353. [PubMed: 8153177]
- Forsgren E. Female sand gobies prefer good fathers over dominant males. *Proc R Soc Lond B Biol Sci*. 1997; 264:1283–1286.
- Fuxjager MJ, Mast G, Becker EA, Marler CA. The ‘home advantage’ is necessary for a full winner effect and changes in post-encounter testosterone. *Horm Behav*. 2009; 56:214–219. [PubMed: 19426733]
- Fuxjager MJ, Forbes-Lorman RM, Coss DJ, Auger CJ, Auger AP, Marler CA. Winning territorial disputes selectively enhances androgen sensitivity in neural pathways related to motivation and social aggression. *Proc Natl Acad Sci USA*. 2010; 107:12393–12398. [PubMed: 20616093]
- Gábor G, Sasser RG, Kastelic JP, Coulter GH, Everson DO, Falkay G, Mézes M, Cook RB, Csik JV, Bárány I, Szász FJ. Effects of GnRH treatment and prediction of sperm output and viability. *Theriogenology*. 1998; 50:177–183. [PubMed: 10734484]
- Gleason ED, Marler CA. Testosterone response to courtship predicts future paternal behavior in the California mouse, *Peromyscus californicus*. *Horm Behav*. 2010; 57:147–154. [PubMed: 19833131]
- Gleason ED, Fuxjager MJ, Oyegbile TO, Marler CA. Testosterone release and social context: when it occurs and why. *Front Neuroendocrinol*. 2009; 30:460–469. [PubMed: 19422843]
- Gottreich A, Zuri I, Barel S, Hammer I, Terkel J. Urinary testosterone levels in the male blind mole rat (*Spalax ehrenbergi*) affect female preference. *Physiol Behav*. 2000; 69:309–315. [PubMed: 10869597]
- Gubernick DJ, Nordby JC. Mechanisms of sexual fidelity in the monogamous California mouse, *Peromyscus californicus*. *Behav Ecol Sociobiol*. 1993; 32:211–219.
- Gubernick DJ, Teferi T. Adaptive significance of male parental care in a monogamous mammal. *Proc R Soc Lond B Biol Sci*. 2000; 267:147–150.
- Jawor JM, McGlothlin JW, Casto JM, Grieves TJ, Snajdr EA, Bentley GE, Ketterson ED. Seasonal and individual variation in response to GnRH challenge in male dark-eyed juncos (*Junco hyemalis*). *Gen Comp Endocrinol*. 2006; 149:182–189. [PubMed: 16814785]
- Ketterson, ED., Parker, KJ., Raouf, SA., Nolan, VJ., Ziegenfuss, C., Chandler, CR. The relative impact of extra-pair fertilizations on variation in male and female reproductive success in dark-eyed juncos (*Junco hyemalis*). In: Parker, PG., Burley, NT., editors. *Avian Reproductive Tactics: Female and Male Perspective*. Allen; Lawrence, KS: 1997.
- Lacombe D, Cyr A, Matton P. Plasma LH and androgen levels in the red-winged blackbird (*Agelaius phoeniceus*) treated with a potent GnRH analogue. *Comp Biochem Physiol, A: Mol Integr Physiol*. 1991; 99A:603–607.

- Lindström K, St Mary CM, Pampoulie C. Sexual selection for male paternal care in the sand goby, *Pomatoschistus minutus*. Behav Ecol Sociobiol. 2006; 60:46–51.
- McGlothlin JW, Parker PG, Nolan VJ, Ketterson ED. Correlational selections leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. Evolution. 2005; 59:658–671. [PubMed: 15856707]
- McGlothlin JW, Jawor JM, Ketterson ED. Natural variation in a testosterone-mediated trade-off between mating effort and parental effort. Am Nat. 2007; 170:864–875. [PubMed: 18171169]
- McGlothlin JW, Jawor JM, Greives TJ, Casto JM, Phillips JL, Ketterson ED. Hormones and honest signals: males with larger ornaments elevate testosterone more when challenged. J Evol Biol. 2008; 21:39–48. [PubMed: 18034801]
- Millesi E, Hoffman IE, Steurer S, Metwaly M, Dittami JP. Vernal changes in the behavioral and endocrine responses to GnRH application in male European ground squirrels. Horm Behav. 2002; 41:51–58. [PubMed: 11863383]
- Osborne J. Notes on the use of data transformations. Practical Assessment, Research and Evaluation. 2002:8.
- Ostlund S, Ahnesjö I. Female fifteen-spined sticklebacks prefer better fathers. Anim Behav. 1998; 56:1177–1183. [PubMed: 9819334]
- Oyegbile TO, Marler CA. Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. Horm Behav. 2005; 48:259–267. [PubMed: 15979073]
- Raouf SA, Parker PG, Ketterson ED, Nolan VJ, Ziegenfuss C. Testosterone affects reproductive success by influencing extra-pair fertilizations in male dark-eyed juncos (Aves: *Junco hyemalis*). Proc R Soc Lond B Biol Sci. 1997; 264:1599–1603.
- Ribble DO. The monogamous mating system of *Peromyscus californicus* as revealed by DNA fingerprint. Behav Ecol Sociobiol. 1991; 29:161–166.
- Ribble DO. Lifetime reproductive success and its correlates in the monogamous rodent, *Peromyscus californicus*. J Anim Ecol. 1992; 61:457–468.
- Ryan KK, Altmann J. Selection for male choice based primarily on mate compatibility in the oldfield mouse, *Peromyscus polionotus rhoadsi*. Behav Ecol Sociobiol. 2001; 50:436–440.
- Spoon TR, Millam JR, Owings DH. Variation in the stability of cockatiel (*Nymphicus hollandicus*) pair relationships: the roles of males, females, and mate compatibility. Anim Behav. 2004; 141:1211–1234.
- Spoon TR, Millam JR, Owings DH. The importance of mate behavioural compatibility in parenting and reproductive success by cockatiels, *Nymphicus hollandicus*. Anim Behav. 2006; 71:315–326.
- Taylor GT, Haller J, Regan D. Female rats prefer an area vacated by a high testosterone male. Physiol Behav. 1982; 28:953–958. [PubMed: 7202224]
- Taylor GT, Regan D, Haller J. Sexual experience, androgens and female choice of a mate in laboratory rats. J Endocrinol. 1983; 96:43–52. [PubMed: 6681625]
- Trainor BC, Marler CA. Testosterone, paternal behavior, and aggression in the monogamous California mouse (*Peromyscus californicus*). Horm Behav. 2001; 40:32–42. [PubMed: 11467882]
- Trainor BC, Marler CA. Testosterone promotes paternal behaviour in a monogamous mammal via conversion to oestrogen. Proc R Soc Lond B Biol Sci. 2002; 269:823–829.
- Trainor BC, Bird IM, Marler CA. Opposing hormonal mechanisms of aggression revealed through short-lived testosterone manipulations and multiple winning experiences. Horm Behav. 2004; 45:115–121. [PubMed: 15019798]
- Trivers, RL. Parental investment and sexual selection. In: Campbell, B., editor. Sexual Selection and the Descent of Man. Aldine; Chicago, IL: 1972. p. 1871-1971.p. 136-179.
- Vickers AJ. The use of percentage from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. BMC Med Res Methodol. 2001; 1:6. [PubMed: 11459516]
- von Schantz T, Göransson G, Andersson G, Fröberg I, Grahn M, Helgée A, Wittzell H. Female choice selects for a viability-based male trait in pheasants. Nature. 1989; 337:166–169. [PubMed: 2911350]
- Wagner WE. Measuring female mating preferences. Anim Behav. 1998; 55:1029–1042. [PubMed: 9632487]

- Yamazaki K, Boyse EA, Mike V, Thaler HT, Mathieson BJ, Abbott J, Boyse J, Zayas ZA, Thomas L. Control of mating preferences in mice by genes in the major histocompatibility complex. *J Exp Med*. 1976; 144:1324–1335. [PubMed: 1032893]
- Ziegler TE, Schultz-Darken NJ, Scott JJ, Snowdon CT, Ferris CF. Neuroendocrine response to female ovulatory odors depends upon social condition in male common marmosets, *Callithrix jacchus*. *Horm Behav*. 2005; 47:56–64. [PubMed: 15579266]

Fig.1.

A 50-ng injection of GnRH causes a significant rise in plasma T as compared to saline injection.

Fig. 2.
Natural log of male baseline T on the day before preference testing plotted against the duration of time a female spent investigating the male.

Fig. 3.

The mean number of days needed for a pair to produce a litter is significantly higher when choosing females are paired with their non-preferred males. Bars represent ± 1 SEM.

Fig. 4.

Bar graph of mean 68-day RS scores. In the left panel, RS scores for pairings of choosing females with their non-preferred and preferred males. On the right, RS scores for naive, non-choosing females on the basis of whether their mate was non-preferred or preferred by another female. Bars represent ± 1 SEM.

Chart 1.
A schematic representation of the experimental design.